Making Informed Decisions Using Data

Where Are We?

- ✓ Decide on a problem to solve.
- ✓ Determine how to measure process/improvement.
- ✓ Understand the current state.
- ❑ Obtain baseline measures
  - ❑ Formulate a data collection plan
  - ❑ Validate the data
- ❑ Analyze the current state and the measures to
  - ❑ Validate assumptions about the process
  - ❑ Prioritize and focus your improvements
  - ❑ Estimate the likely results of your changes

Essentials of Data Collection in QI

Quality Improvement

- Formulate your goal
- Define your scope
- Identify measures
  - Elements
  - Operational Definitions
  - Units of Measure
- Use available data
- Leverage appropriate statistical tools for analysis
- P-Value is interesting

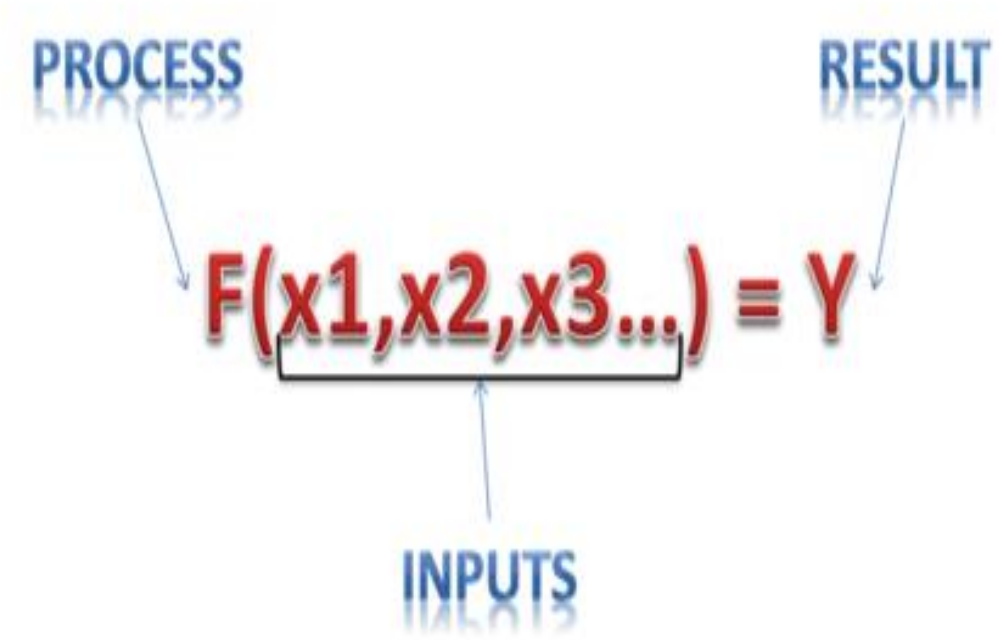
Research

- Formulate your question
- Define your population
- Define your methods
- Carefully control your data sources
- Leverage appropriate statistical tools for analysis
- P value is king

Look to the Outcome

Output/Outcome is a product of the components or processes that produce it.

$$Y = f(x)$$



http://media.tumblr.com/f592067382572a403457dbe86a150fee/tumblr_inline_mjvfr0n4SQ1qz4rgp.png

The Basic Quality Equation

$Y=f(x)$

Commute time = (time of day) + (driving speed) + (# of traffic lights) + (route chosen)

Essentials of Data Collection Plans in QI

1. Determine the questions you want to answer about your process
2. Think one level deeper than the surface question- segmentation factors
3. List out your data elements
4. Define each with an operational definition- equation, units of measure, location of data element in Epic
5. Identify what this data element is needed for – Y measure or one of the X's
6. What is the defect definition for this element

Now is the time to exercise restraint!

Data Collection Plan

D M A I C

Data Element Name	Y/X	Type	Operational Definition	Source	Unit of measure	Defect Defn
	Y					
	X					
	X					
	X					
	X					
	X					
	X					

Key Questions: (ex: how much, how many, by whom)

- A
- B
- C
- D.

Segmentation factors: (ex: provider name, clinic name, day of week etc…)

- A
- B
- C
- D.

Let’s formulate a data collection plan for your project.

- In table teams, use the data collection worksheet to list what data elements you would ask for.
  - What are the questions you are trying to answer?
    - How often does something happen?
    - How long does each part of the process take to complete?
    - What takes the most time?
  - If you had answers to these what other information might you suddenly want?
    - What are some key segmentation factors you would want to get to help answer these additional questions?

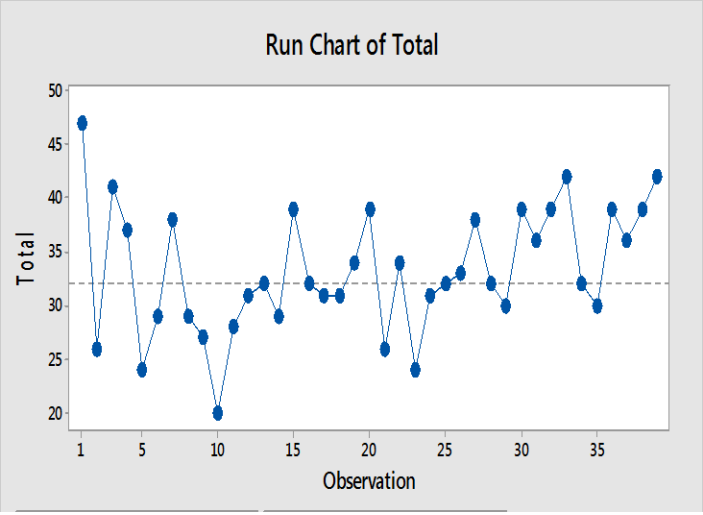
Data Collection Tip

Keep your data request as simple as possible to avoid delays.

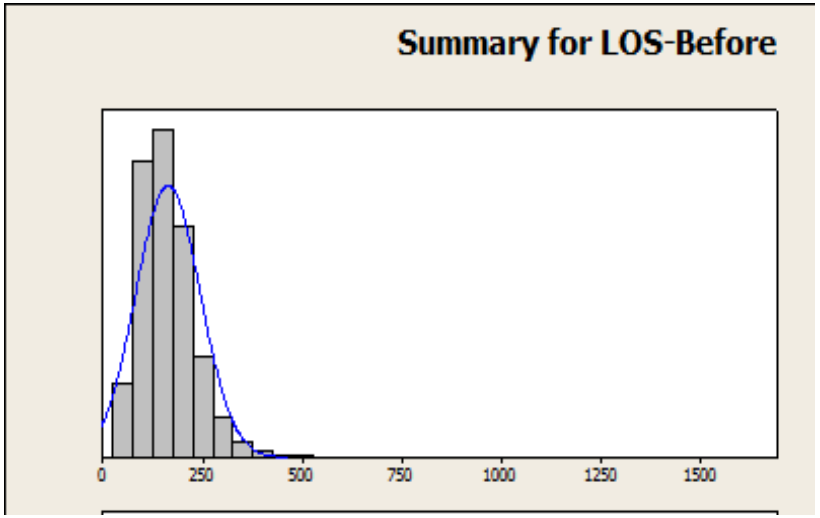
Using Data Once You Get It

Two Views

- Look at data with respect to time



- Look at the shape, center and spread of the data



Using and Interpreting Data Effectively

Say…what's a mountain goat doing way up here in a cloud bank?

Sources of Variation

Variation is inherent in everything. The key is whether or not it is predictable variation or unpredictable variation.

Two sources of variation:

1. Common Cause- variation inherent in the process
2. Special Cause- variation from outside the normal process

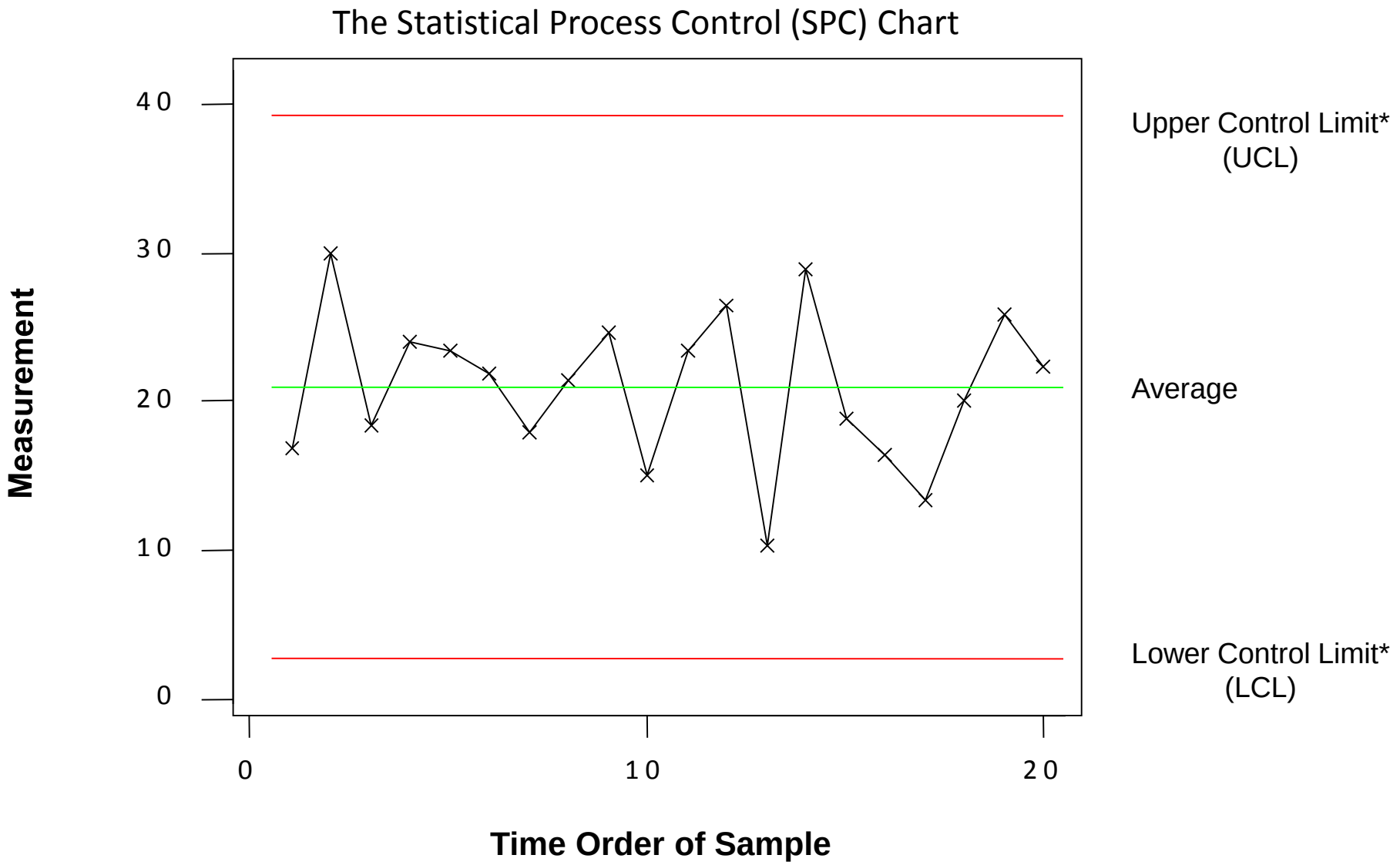
Two Possible Mistakes:

Mistake 1: To react to an outcome as if it came from a special cause, when actually it came from common causes of variation.

Mistake 2: To react to an outcome as if it came from common causes of variation, when actually it came from a special cause.

SPC charts facilitate correct identification of variation sources.

Looking at Data With Respect to Time

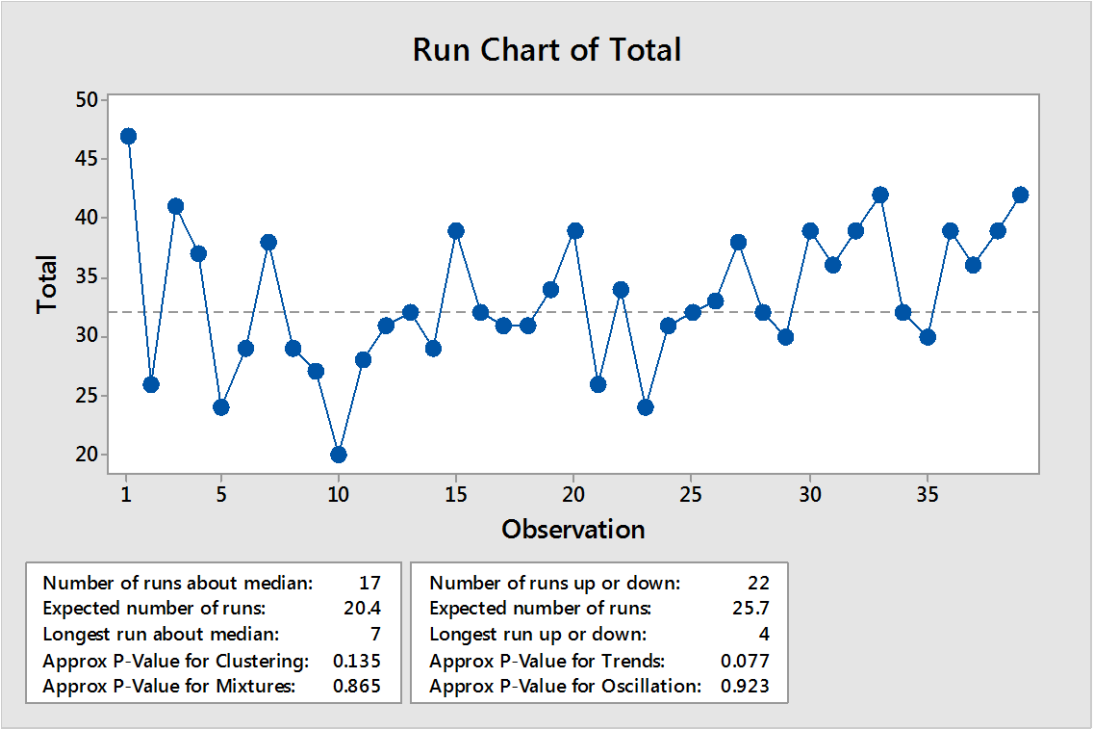


Control Chart Basics

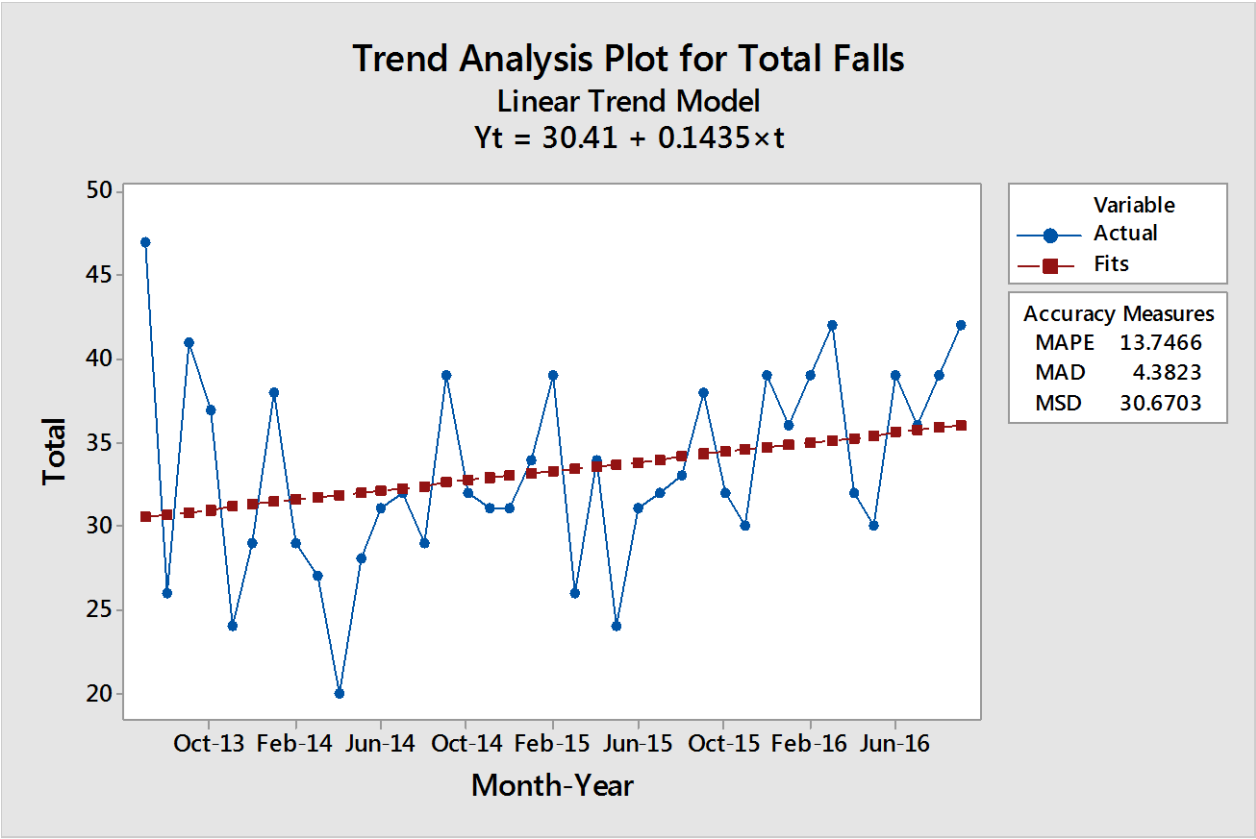
- 3σ-level Control Limits
  - Created by Dr. Walter Shewhart to minimize two types of mistakes
  - Placed empirically because they minimize the two types of mistakes
  - Are not probability limits
- Two types of Mistakes:
  - Calling a special cause of variation a common cause of variation (Missing a chance to identify a change in the process)
  - Calling a common cause of variation a special cause of variation (Interfering with a stable process or tampering, wasting resources looking for special causes of variation that do not exist)
- 8 Tests of Variation
  - Empirically derived by Shewhart
  - Common patterns to look for that indicate special cause variation

We want best efforts guided by theory. Confusing common causes with special causes will only make things worse. *W. Edwards Deming*

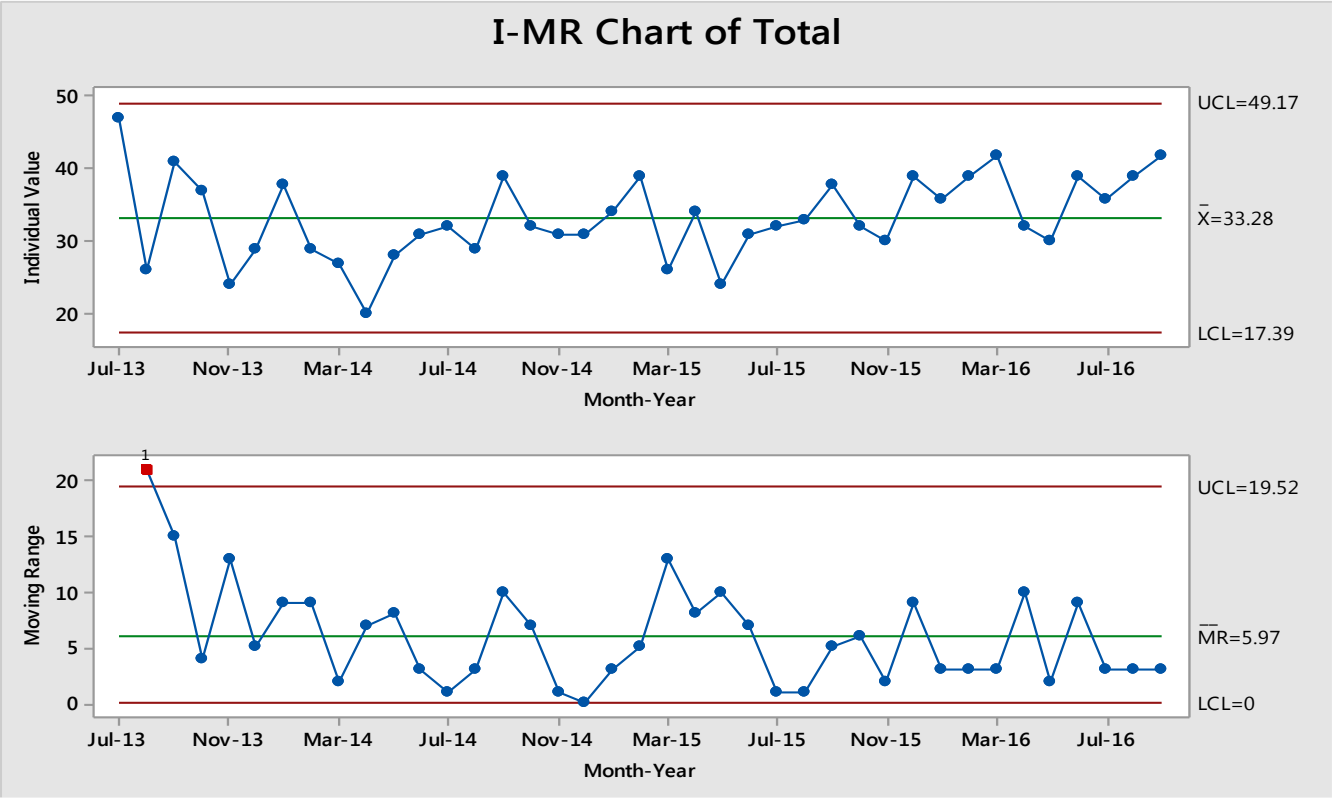
Run Charts



- Run charts show variation with respect to time. Run charts should look random. Non random behavior is an indication of special cause variation. Variation that is due to something other than chance or common cause variation

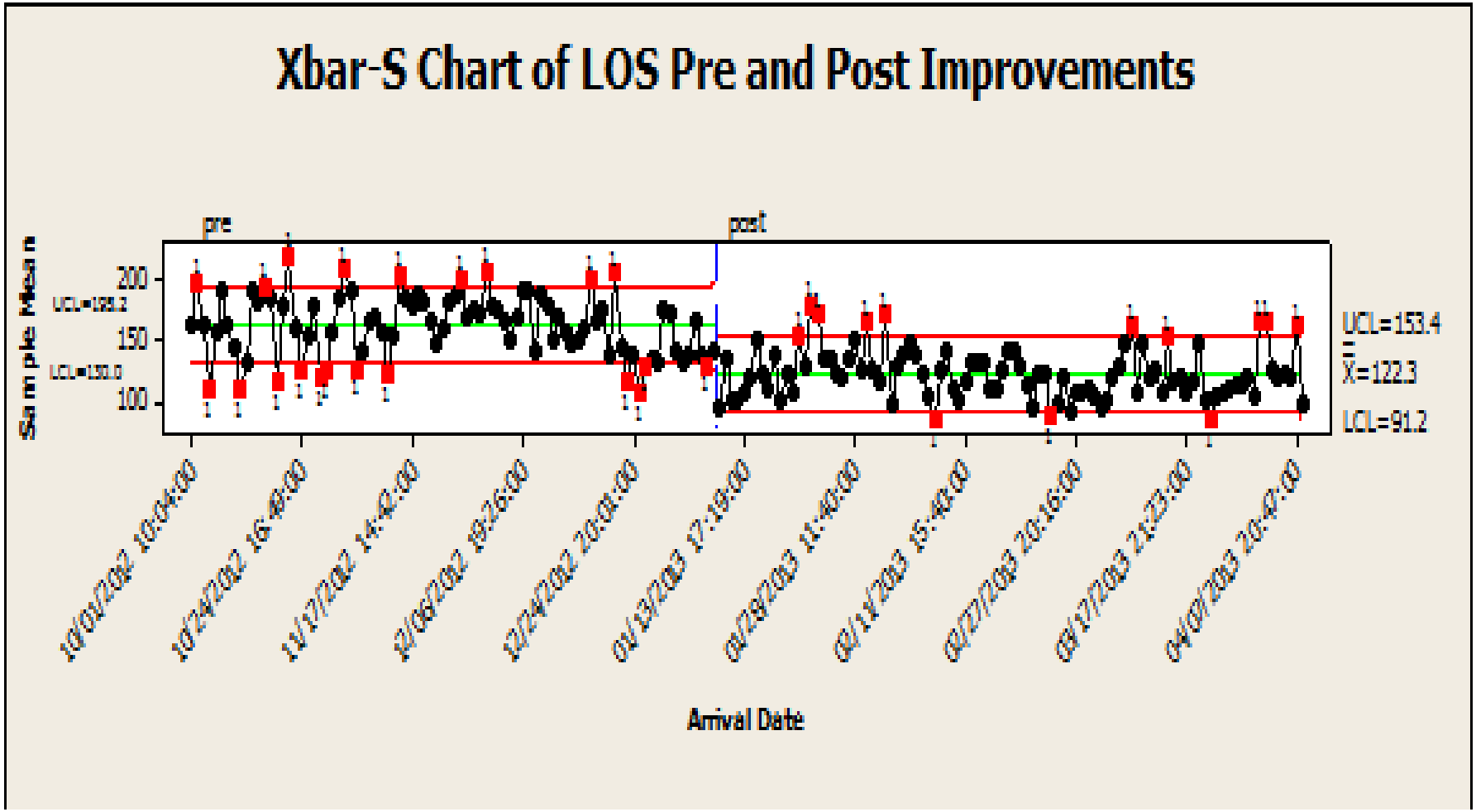


Are we getting worse?



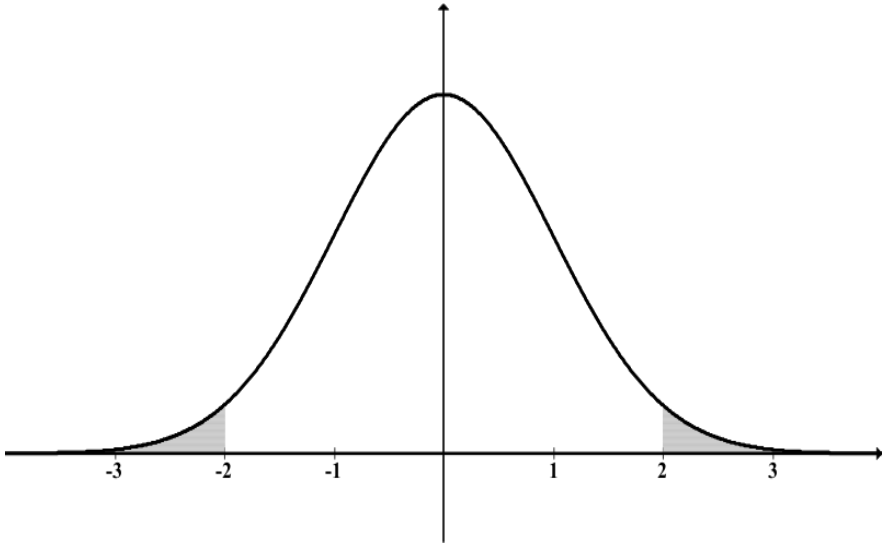
Actually we appear to have a stable system that produces a predictable number of falls month after month since 2013.

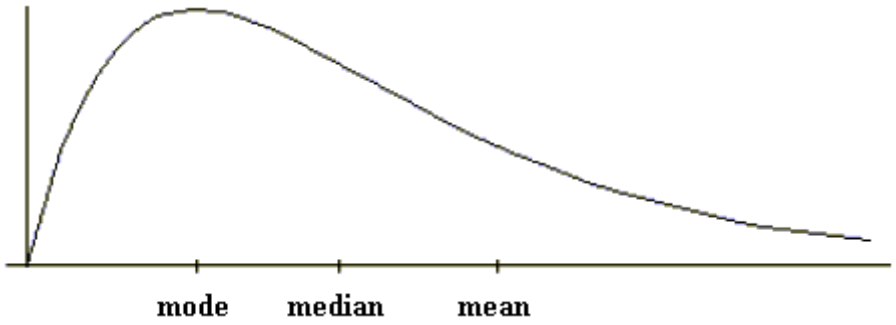
Did We Improve?



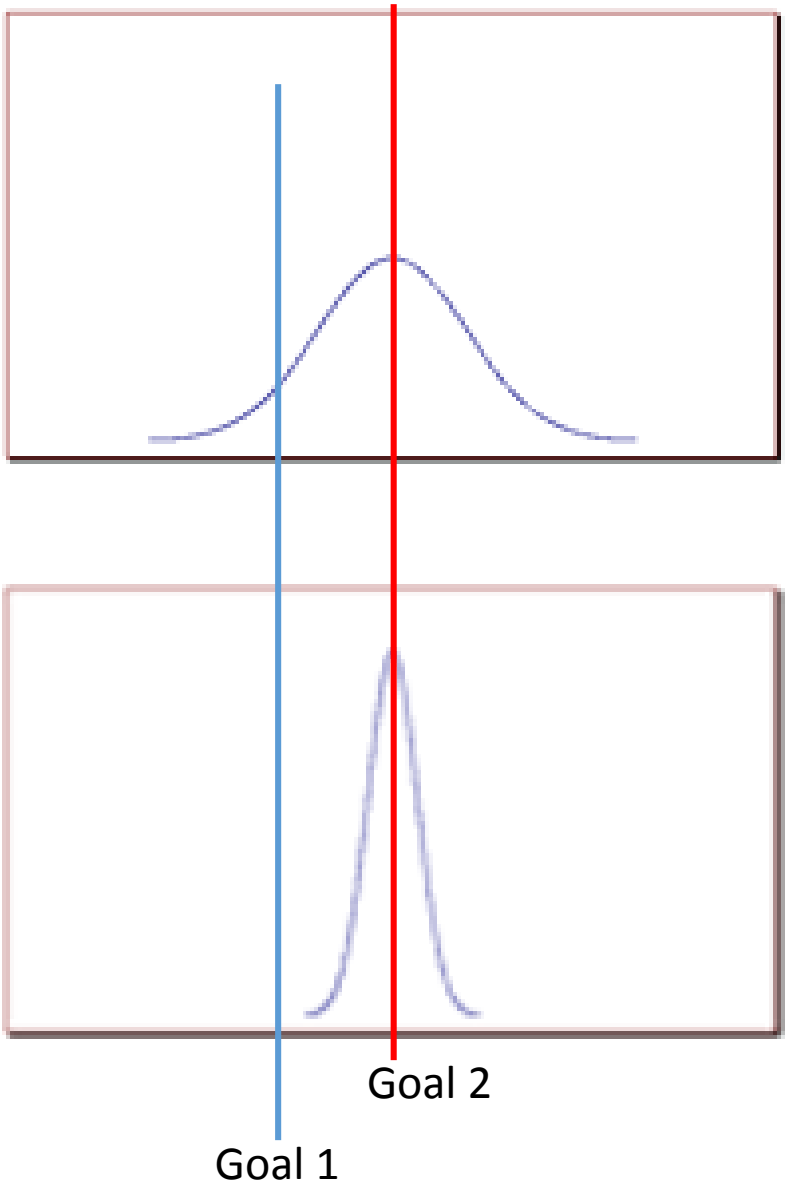
Looking at Shape, Center, Spread of Data

Shape Matters

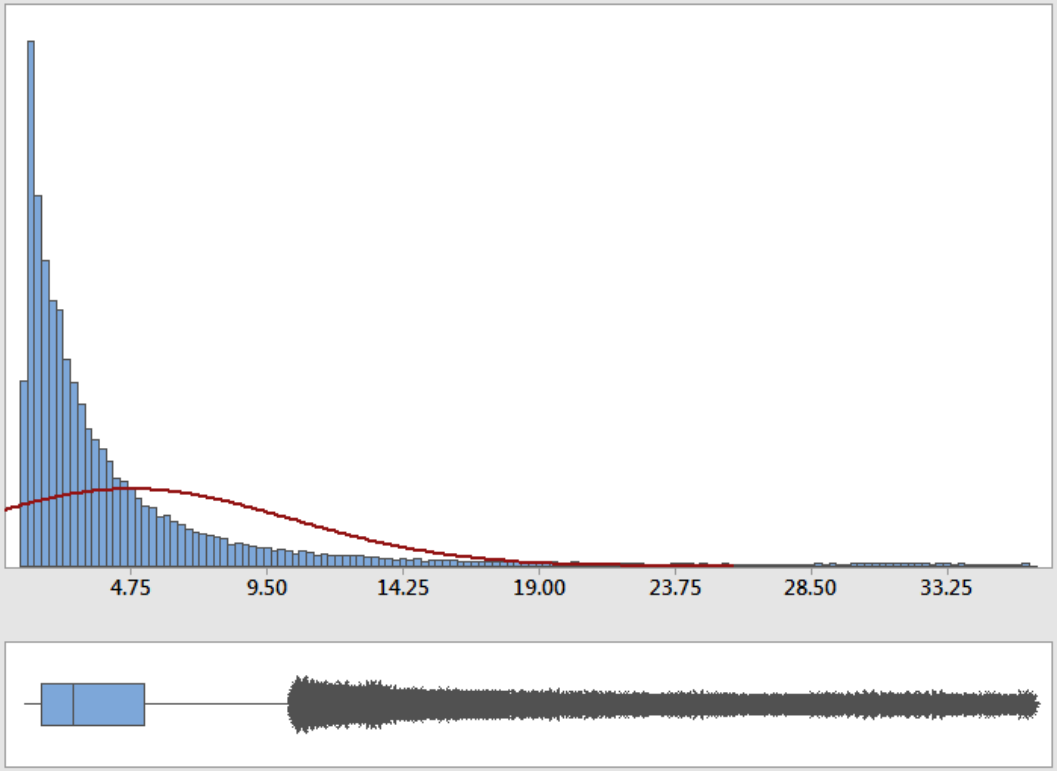




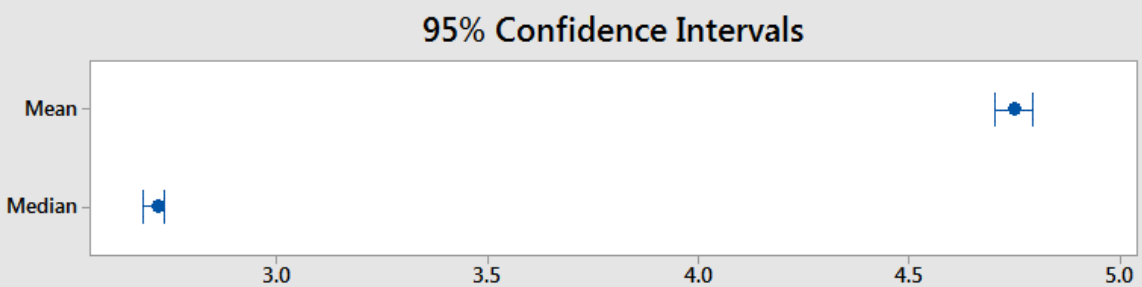
Variation Matters



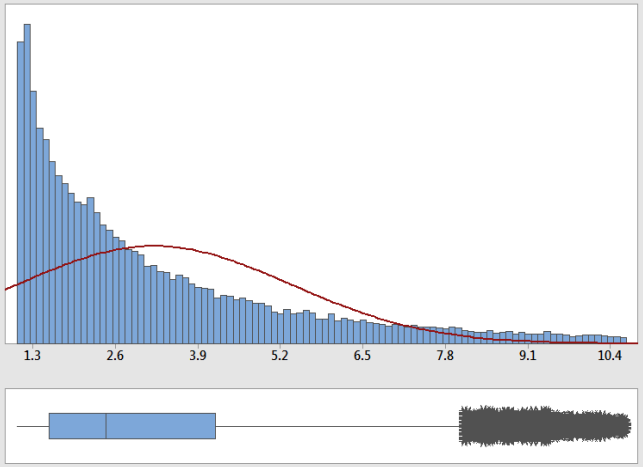
Summary Report for Incoming Call Duration



Anderson-Darling Normality Test	
A-Squared	7466.01
P-Value	<0.005
Mean	4.7503
StDev	5.6599
Variance	32.0340
Skewness	3.0251
Kurtosis	10.2193
N	63167
Minimum	1.0500
1st Quartile	1.6333
Median	2.7167
3rd Quartile	5.1833
Maximum	36.2167
95% Confidence Interval for Mean	
4.7061	4.7944
95% Confidence Interval for Median	
2.6833	2.7333
95% Confidence Interval for StDev	
5.6288	5.6912

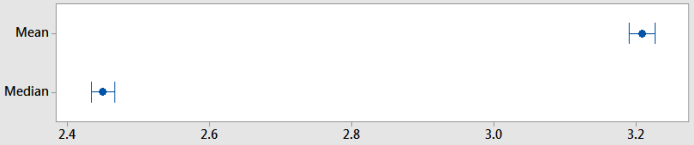


Summary Report fo Incoming Call Duration, Trimmed

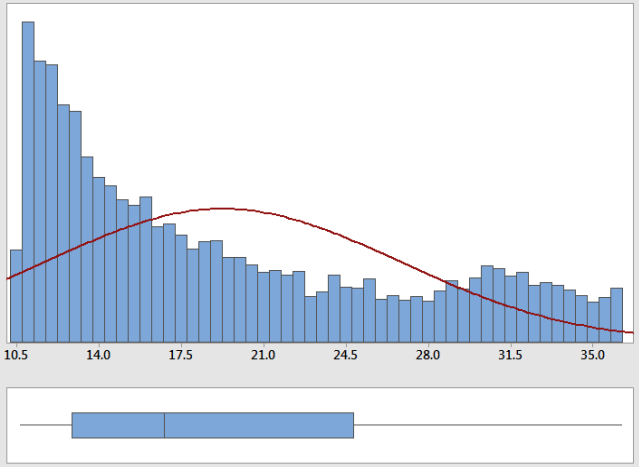


Anderson-Darling Normality Test	
A-Squared	2838.86
P-Value	<0.005
Mean	3.2084
StDev	2.1619
Variance	4.6739
Skewness	1.35580
Kurtosis	1.21130
N	57074
Minimum	1.0500
1st Quartile	1.5500
Median	2.4500
3rd Quartile	4.1667
Maximum	10.6333
95% Confidence Interval for Mean	
3.1907	3.2261
95% Confidence Interval for Median	
2.4333	2.4667
95% Confidence Interval for StDev	
2.1494	2.1745

95% Confidence Intervals

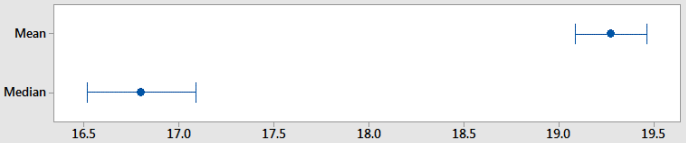


Summary Report for Incoming Call Duration, Outlier



Anderson-Darling Normality Test	
A-Squared	234.40
P-Value	<0.005
Mean	19.275
StDev	7.504
Variance	56.315
Skewness	0.722508
Kurtosis	-0.776300
N	6063
Minimum	10.650
1st Quartile	12.850
Median	16.800
3rd Quartile	24.817
Maximum	36.217
95% Confidence Interval for Mean	
19.086	19.464
95% Confidence Interval for Median	
16.517	17.088
95% Confidence Interval for StDev	
7.373	7.640

95% Confidence Intervals



Key Learnings

- Determine what questions you want to answer before requesting data
  - Only ask for what you need- one puppy!
  - Think through what other questions you might have once you know the answers to your main questions
- Validate your data
- Understand the factors that contribute to variation in the data
- Use judgment to interpret the data- you haven’t controlled for major variables in a sample of convenience
- Let the data guide and inform your focused improvements